

Álgebra linear e compressão de imagens JPEG

Prof. Miguel Frasson

Depto. de Matemática Aplicada e Estatística – ICMC
USP São Carlos

Resumo

- ▶ Apresentamos a compressão de imagens no formato JPEG.
- ▶ Ingrediente principal: **Transformada Discreta de Cosseno**.

Transformada Discreta de Cosseno (TDC)

- ▶ mudança de bases em $M_8(\mathbb{R})$ para uma nova base ortogonal
- ▶ nessa nova base, a informação essencial é mais evidente

Mudança de base

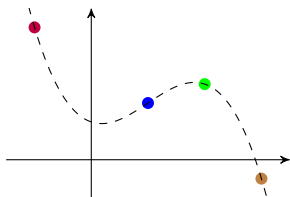
- ▶ Na **Álgebra Linear**, organizamos informação através de **coordenadas**.
 - ▶ $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ base de um espaço vetorial V
 - ▶ $\forall v \in V: v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$
- ▶ Muitas vezes, as informações podem ser mais facilmente visualizadas se mudamos o ponto de vista.
- ▶ Isso pode ser feito pela escolha criteriosa de uma base.

Exemplo — interpolação

Problema

Determinar um polinômio $p(x)$ de grau mínimo tal que

$$p(-1) = 7, \quad p(1) = 3, \quad p(2) = 4, \quad p(3) = -1.$$

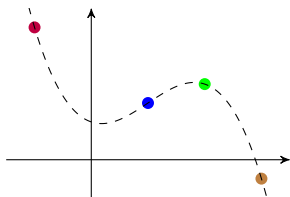


Exemplo — interpolação

Problema

Determinar um polinômio $p(x)$ de grau mínimo tal que

$$p(-1) = 7, \quad p(1) = 3, \quad p(2) = 4, \quad p(3) = -1.$$



$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 ?$$

$\mathcal{P}_3(\mathbb{R})$: base canônica

$$p(-1) = 7, \quad p(1) = 3, \quad p(2) = 4, \quad p(3) = -1$$

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$$

Base canônica

- ▶ $1, x, x^2, x^3$
- ▶ Resolver esse problema = resolver um sistema 4×4

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \\ 4 \\ -1 \end{pmatrix}$$

- ▶ $\therefore p(x) = 2 - x + 3x^2 - x^3$

$\mathcal{P}_3(\mathbb{R})$: polinômios de Legendre

$$p(-1) = 7, \quad p(1) = 3, \quad p(2) = 4, \quad p(3) = -1$$

- Polinômios de Legendre L_n em $A = \{-1, 1, 2, 3\}$

n	$L_n(-1)$	$L_n(1)$	$L_n(2)$	$L_n(3)$
-1	1	0	0	0
1	0	1	0	0
2	0	0	1	0
3	0	0	0	1

- Exemplo: $L_1(x) = \frac{(x+1)(x-2)(x-3)}{4}$

- $\therefore p(x) = 7L_{-1}(x) + 3L_1(x) + 4L_2(x) - L_3(x)$

Produto interno

Produto interno

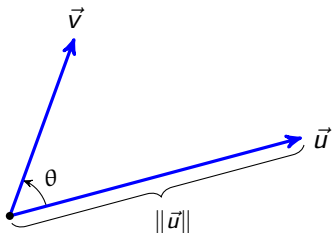
$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

- ▶ $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- ▶ $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle, \alpha \in \mathbb{R}$
- ▶ $\langle x, y \rangle = \langle y, x \rangle$
- ▶ $\langle x, x \rangle > 0$ para $x \neq 0$

Produto interno

Produto interno: ferramenta importante para:

- ▶ Medir comprimentos e distâncias: $\|x\| = \sqrt{\langle x, x \rangle}$
- ▶ Ângulos retos: $x \perp y \iff \langle x, y \rangle = 0$
- ▶ Medir ângulos: $\cos \theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$



Bases ortogonais

- ▶ Quando $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ é **base ortogonal** ($u_i \perp u_j, i \neq j$) é fácil calcular coordenadas de

$$v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$$

na base \mathcal{B} :

$$\begin{aligned}\langle v, u_1 \rangle &= \langle \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n, u_1 \rangle \\ &= \alpha_1 \langle u_1, u_1 \rangle + \alpha_2 \langle u_2, u_1 \rangle + \dots + \alpha_n \langle u_n, u_1 \rangle \\ &= \alpha_1 \|u_1\|^2\end{aligned}$$

(Note: In the original image, arrows point from the terms $\alpha_2 \langle u_2, u_1 \rangle$ and $\alpha_n \langle u_n, u_1 \rangle$ to a '0', indicating they are zero due to orthogonality.)

- ▶ $\therefore \alpha_i = \frac{\langle v, u_i \rangle}{\|u_i\|^2}, \forall i = 1, \dots, n$
- ▶ Além disso, \mathcal{B} ortogonal nos dá de maneira fácil a mudança de base de retorno à base original.

JPEG

- ▶ Sigla para **Joint Photographic Experts Group**.
- ▶ **1992**: primeiro rascunho do padrão.

Princípio

Estatisticamente, imagens de câmeras digitais têm variações suaves de cores entre *pixels* vizinhos.

Principal ingrediente

A aplicação da Transformada Discreta de Cosseno, para filtrar tais variações

Transformada Discreta de Cosseno (TDC)

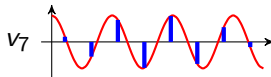
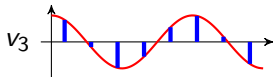
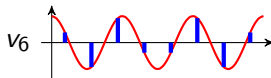
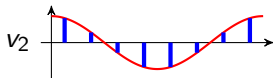
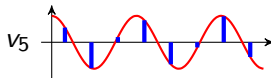
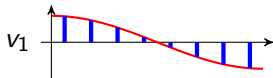
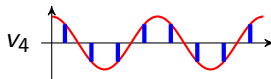
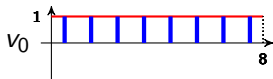
- ▶ Ahmed, Natarajan, Rao (1974)
- ▶ $TDC : \mathbb{R}^n \rightarrow \mathbb{R}^n$
- ▶ Lembra a Transformada de Fourier \rightarrow **espaço \times frequência**
- ▶ Pode ser vista como uma mudança de base, na qual as “frequências” estão evidentes.

Uso da TDC no JPEG

Dados reais: **transição suave** de valores para entradas vizinhas
 \rightarrow coeficientes de “altas frequências” desprezíveis
 \rightarrow truncamento \rightarrow compressão com perdas.

TDC 1-D

- ▶ frequência n em $\{0, 1, 2, \dots, 7\}$
- ▶ v_n o vetor que interpola a função $\cos n\frac{\pi}{8}t$ em $t = i + \frac{1}{2}$:



Teorema: v_n é uma base ortogonal de \mathbb{R}^8 .

Demonstração.

$$\blacktriangleright \sum_{i=0}^7 \cos \frac{\pi}{8} \left(i + \frac{1}{2} \right) = \begin{cases} 8, & k = 0, \\ 0, & 0 < k < 16. \end{cases}$$

Teorema: v_n é uma base ortogonal de \mathbb{R}^8 .

Demonstração.

$$\blacktriangleright \sum_{i=0}^7 \cos \frac{\pi}{8} \left(i + \frac{1}{2} \right) = \begin{cases} 8, & k = 0, \\ 0, & 0 < k < 16. \end{cases}$$

$$\sin(a + b) - \sin(a - b) = 2 \sin b \cos a$$

$$\sin(a + b) = \sin a \cos b + \sin b \cos a$$

$$\sin(a - b) = \sin a \cos b - \sin b \cos a \quad -$$

$$\hline \sin(a + b) - \sin(a - b) = 2 \sin b \cos a$$

Teorema: v_n é uma base ortogonal de \mathbb{R}^8 .

Demonstração.

$$\blacktriangleright \sum_{i=0}^7 \cos k \frac{\pi}{8} \left(i + \frac{1}{2}\right) = \begin{cases} 8, & k = 0, \\ 0, & 0 < k < 16. \end{cases}$$

$$\sin(a + b) - \sin(a - b) = 2 \sin b \cos a$$

$$C = \sum_{i=0}^7 \cos k \frac{\pi}{8} \left(i + \frac{1}{2}\right)$$

$$\begin{aligned} \overbrace{2 \sin \left(k \frac{\pi}{8} \frac{1}{2}\right)}^{\neq 0, 0 < k < 16} \times C &= \sum_{i=0}^7 \overbrace{2 \sin \left(k \frac{\pi}{8} \frac{1}{2}\right)}^b \times \overbrace{\cos \left(k \frac{\pi}{8} \left(i + \frac{1}{2}\right)\right)}^a \\ &= \sum_{i=0}^7 \left[\sin \left(k \frac{\pi}{8} (i + 1)\right) - \sin \left(k \frac{\pi}{8} i\right) \right] \\ &= \sin k\pi - \sin 0 = 0 \quad \therefore \quad C = 0 \end{aligned}$$

Teorema: v_n é uma base ortogonal de \mathbb{R}^8 .

Demonstração.

$$\blacktriangleright \sum_{i=0}^7 \cos \frac{\pi}{8} \left(i + \frac{1}{2}\right) = \begin{cases} 8, & k = 0, \\ 0, & 0 < k < 16. \end{cases}$$

$$\begin{aligned} \blacktriangleright \cos a \times \cos b &= \frac{\cos(a+b)}{2} + \frac{\cos(a-b)}{2} \\ \cos(a+b) &= \cos a \cos b - \sin a \sin b \\ \cos(a-b) &= \cos a \cos b + \sin a \sin b \\ \hline \cos(a+b) + \cos(a-b) &= 2 \cos b \cos a \end{aligned}$$

Teorema: v_n é uma base ortogonal de \mathbb{R}^8 .

Demonstração.

$$\blacktriangleright \sum_{i=0}^7 \cos \frac{\pi}{8} \left(i + \frac{1}{2} \right) = \begin{cases} 8, & k = 0, \\ 0, & 0 < k < 16. \end{cases}$$

$$\blacktriangleright \cos a \times \cos b = \frac{\cos(a+b)}{2} + \frac{\cos(a-b)}{2}$$

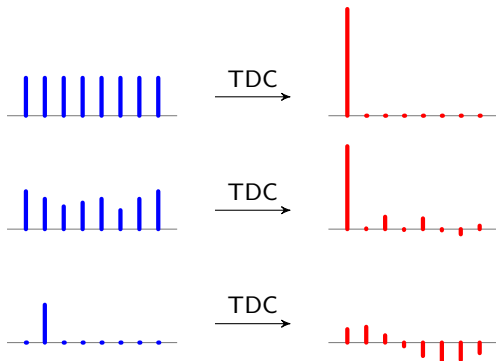
$$\blacktriangleright n \neq m:$$

$$\begin{aligned} \langle v_n, v_m \rangle &= \sum_{t=0}^7 \cos \frac{n\pi}{8} \left(t + \frac{1}{2} \right) \times \cos \frac{m\pi}{8} \left(t + \frac{1}{2} \right) \\ &= \frac{1}{2} \sum_{t=0}^7 \cos \frac{(n+m)\pi}{8} \left(t + \frac{1}{2} \right) \\ &\quad + \frac{1}{2} \sum_{t=0}^7 \cos \frac{(n-m)\pi}{8} \left(t + \frac{1}{2} \right) = 0. \quad \square \end{aligned}$$

TDC 1-D

Transformada Discreta de Cosseno (TDC) 1-D

$\text{TDC}(v) : \mathbb{R}^8 \rightarrow \mathbb{R}^8$ é a **mudança de base** da base canônica para a base ortonormal v_0, v_1, \dots, v_7 .

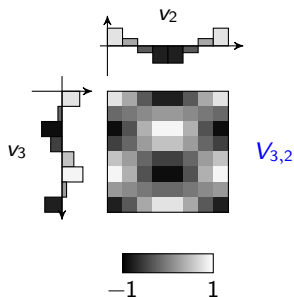


TDC 2-D

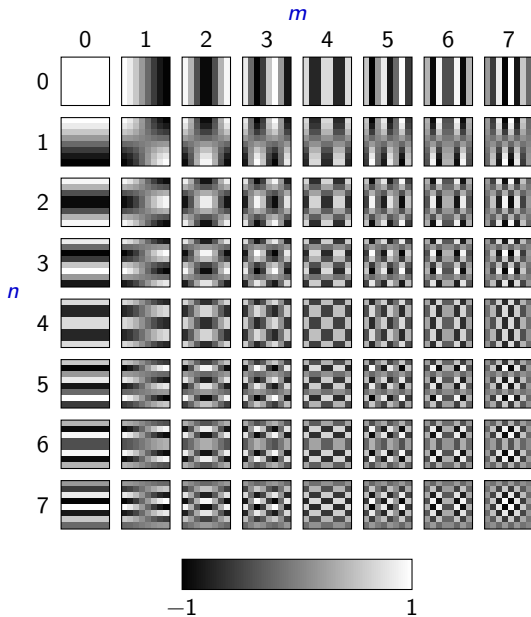
- Podemos construir a TDC 2-D a partir da TDC 1-D.

- $V_{n,m} = (a_{n,m}(i,j)), \quad 0 \leq n, m < 8$

$$a_{n,m}(i,j) = \cos \frac{n\pi}{8} \left(i + \frac{1}{2}\right) \times \cos \frac{m\pi}{8} \left(j + \frac{1}{2}\right)$$



TDC 2-D: os 64 elementos $V_{n,m}$



Teorema: $V_{n,m}$ é ortogonal em $M_{8 \times 8}(\mathbb{R})$

Demonstração.

Suponha $(n, m) \neq (k, \ell)$:

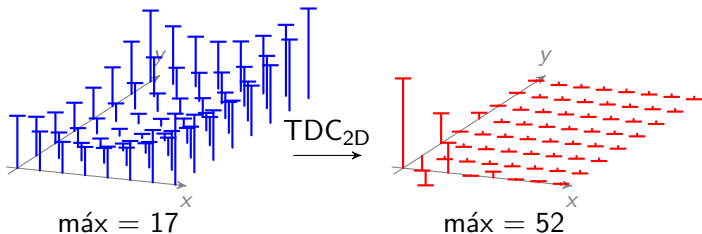
$$\begin{aligned}\langle V_{n,m}, V_{k,\ell} \rangle &= \sum_{i=0}^7 \sum_{j=0}^7 \cos \frac{n\pi}{8} \left(i + \frac{1}{2}\right) \times \cos \frac{m\pi}{8} \left(j + \frac{1}{2}\right) \\ &\quad \times \cos \frac{k\pi}{8} \left(i + \frac{1}{2}\right) \times \cos \frac{\ell\pi}{8} \left(j + \frac{1}{2}\right) \\ &= \left[\sum_{i=0}^7 \cos \frac{n\pi}{8} \left(i + \frac{1}{2}\right) \times \cos \frac{k\pi}{8} \left(i + \frac{1}{2}\right) \right] \\ &\quad \times \left[\sum_{j=0}^7 \cos \frac{m\pi}{8} \left(j + \frac{1}{2}\right) \times \cos \frac{\ell\pi}{8} \left(j + \frac{1}{2}\right) \right] \\ &= \langle v_n, v_k \rangle_{\mathbb{R}^8} \times \langle v_m, v_\ell \rangle_{\mathbb{R}^8} = 0.\end{aligned}$$

□

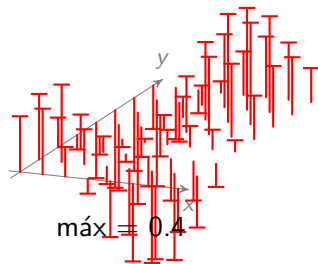
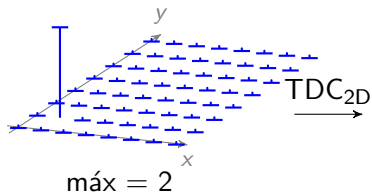
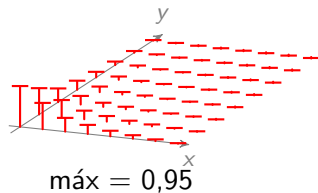
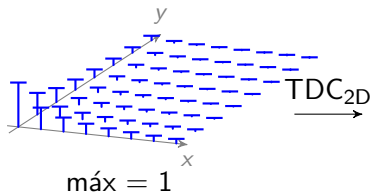
TDC 2-D

Transformada Discreta de Cosseno (TDC) 2-D

$\text{TDC}(v) : M_8(\mathbb{R}) \rightarrow M_8(\mathbb{R})$ é a **mudança de base** da base canônica para a base ortonormal $V_{n,m}$, $n, m = 0, \dots, 7$.



TDC 2-D



JPEG

- ▶ Separar as cores em *brilho* (cinza) + *azul* + *vermelho*.
- ▶ Sensibilidade a brilho é maior que a cores
→ reduzir a resolução das cores à metade.
- ▶ Dividir a imagem em blocos 8×8 : A para cada cor.
- ▶ Aplicar a TDC 2-D: $B = \text{TDC}(A)$ em cada bloco
- ▶ Quantização $C = Q(B)$: diminuir a *resolução* dos números para inteiros distantes
→ boa parte dos coeficientes de B tornam-se 0.
→ escrever a matriz de quantização no início do arquivo JPEG.
- ▶ alinhar as entradas de C em *ziguezague*
→ zeros no final desprezados.
- ▶ comprime esse vetor de inteiros com um algoritmo sem perdas e escreve-o no arquivo JPEG.

Quantização

- ▶ Seja a resolução r
- ▶ Dado $x \in \mathbb{R}$, escolher o número nr ($n \in \mathbb{Z}$) mais próximo de x

$$n = \text{arredondar}\left(\frac{x}{r}\right)$$

Esse processo chama-se **quantização**.

Quantização

- ▶ Seja a resolução r
- ▶ Dado $x \in \mathbb{R}$, escolher o número nr ($n \in \mathbb{Z}$) mais próximo de x

$$n = \text{arredondar}\left(\frac{x}{r}\right)$$

Esse processo chama-se **quantização**.

- ▶ No JPEG, dá-se uma matriz 8×8 de resoluções, chamada de **matriz de quantização Q** , escrita no início do arquivo.
- ▶ Chamamos de $Q(B)$ à matriz dada pela quantização a cada entrada de B .

▶ Exemplo: $Q = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix}.$

Exemplo – bloco 8×8



$$A = \begin{pmatrix} 84 & 83 & 82 & 77 & 77 & 88 & 118 & 152 \\ 93 & 96 & 107 & 123 & 137 & 126 & 100 & 102 \\ 155 & 156 & 181 & 192 & 166 & 92 & 51 & 46 \\ 222 & 182 & 173 & 155 & 115 & 92 & 111 & 137 \\ 226 & 194 & 165 & 189 & 213 & 226 & 232 & 233 \\ 227 & 219 & 218 & 241 & 255 & 255 & 248 & 223 \\ 225 & 215 & 231 & 225 & 213 & 223 & 207 & 185 \\ 152 & 135 & 149 & 156 & 160 & 183 & 196 & 215 \end{pmatrix}$$

Exemplo – passo TDC

$$A = \begin{pmatrix} 84 & 83 & 82 & 77 & 77 & 88 & 118 & 152 \\ 93 & 96 & 107 & 123 & 137 & 126 & 100 & 102 \\ 155 & 156 & 181 & 192 & 166 & 92 & 51 & 46 \\ 222 & 182 & 173 & 155 & 115 & 92 & 111 & 137 \\ 226 & 194 & 165 & 189 & 213 & 226 & 232 & 233 \\ 227 & 219 & 218 & 241 & 255 & 255 & 248 & 223 \\ 225 & 215 & 231 & 225 & 213 & 223 & 207 & 185 \\ 152 & 135 & 149 & 156 & 160 & 183 & 196 & 215 \end{pmatrix}$$

$$B = \text{TDC}(A)$$

$$= \begin{pmatrix} 1313.1 & 21.1 & -6.0 & 7.0 & 29.6 & 9.9 & 6.7 & 0.7 \\ -325.0 & 45.7 & -13.6 & -26.6 & 15.9 & -14.5 & -3.6 & 3.2 \\ -145.2 & -97.8 & 4.1 & -16.7 & -18.3 & -4.0 & 9.7 & -1.3 \\ 138.1 & -98.8 & 37.3 & 29.2 & -17.3 & -1.7 & -9.0 & -3.8 \\ -70.1 & -61.1 & 109.7 & -13.6 & 8.9 & 3.6 & 0.7 & -2.1 \\ -13.2 & 94.6 & 14.0 & -35.0 & -3.0 & 12.2 & 4.4 & 3.7 \\ 0.5 & 7.7 & -26.1 & -11.1 & 9.1 & -2.0 & -6.6 & 3.4 \\ 23.0 & 6.4 & -25.7 & 9.5 & 0.7 & -4.4 & -1.4 & -0.5 \end{pmatrix}$$

Exemplo – passo quantização

$$Q = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix}$$

$$C = Q(B) = \begin{pmatrix} 82 & 2 & -1 & 0 & 1 & 0 & 0 & 0 \\ -27 & 4 & -1 & -1 & 1 & 0 & 0 & 0 \\ -10 & -8 & 0 & -1 & 0 & 0 & 0 & 0 \\ 10 & -6 & 2 & 1 & 0 & 0 & 0 & 0 \\ -4 & -3 & 3 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Uma versão comprimida de C vai escrita no arquivo JPEG.

Exemplo – desfazendo a compressão

\otimes = multiplicação entrada a entrada respectiva

$$D = \text{iTDC}(C \otimes Q) = \begin{pmatrix} 86 & 76 & 76 & 82 & 78 & 80 & 115 & 160 \\ 99 & 92 & 108 & 141 & 152 & 129 & 104 & 96 \\ 152 & 161 & 178 & 181 & 145 & 88 & 53 & 47 \\ 206 & 198 & 184 & 161 & 125 & 98 & 102 & 121 \\ 229 & 191 & 159 & 168 & 203 & 230 & 237 & 233 \\ 233 & 223 & 221 & 237 & 255 & 255 & 248 & 230 \\ 201 & 222 & 240 & 236 & 216 & 197 & 190 & 190 \\ 145 & 145 & 144 & 149 & 166 & 189 & 207 & 214 \end{pmatrix}$$



original



jpeg

Exemplo – comparação



Exemplo – comparação



Referências



Ahmed, N., Natarajan, T., Rao, K. R.

Discrete Cosine Transform.

IEEE Transactions on Computers C-23 (1): 90-93. (1974)

DOI: 10.1109/T-C.1974.223784.



ITU (International Telecommunication Union)

Information Technology – Digital Compression and Coding of Continuous-Tone Still Images – Requirements and Guidelines/Recommendation T.81 (1992)

<http://www.w3.org/Graphics/JPEG/itu-t81.pdf>



JPEG – Idea and Practice

https://en.wikibooks.org/wiki/JPEG_-_Idea_and_Practice